

A Factor-Graph Representation of Probabilities in Quantum Mechanics

Hans-Andrea Loeliger
ETH Zurich
loeliger@isi.ee.ethz.ch

Pascal O. Vontobel
Hewlett-Packard Laboratories, Palo Alto
pascal.vontobel@ieee.org

Abstract—A factor-graph representation of quantum-mechanical probabilities is proposed. Unlike standard statistical models, the proposed representation uses auxiliary variables (state variables) that are not random variables.

I. INTRODUCTION

Statistical models with many variables are often represented by factor graphs [1]–[4] or similar graphical representations [5]–[7]. Such graphical representations can be helpful in various ways, including elucidation of the model itself as well as the derivation of algorithms for statistical inference.

So far, however, quantum mechanics (e.g., [8], [9]) has been standing apart. Despite being a statistical theory, quantum mechanics does not seem to fit into standard statistical categories. Indeed, it has often been emphasized that quantum mechanics is a generalization of probability theory that cannot be understood in terms of “classical” statistical modeling.

In this paper, we propose the different perspective that the probabilities in quantum mechanics are quite ordinary, but their state-space representation is of a type not previously used in statistical modeling. In particular, we propose a factor-graph representation of quantum mechanics that correctly represents the joint probability distribution of any number of measurements. Like most statistical models, the proposed factor graphs use auxiliary variables (state variables) in addition to the actually observed variables; however, in contrast to standard statistical models, the auxiliary variables in the proposed factor graphs are not random variables. Nonetheless, the probabilities of the observations are marginals of the factor graph, as in standard statistical models.

The paper is structured as follows. Section II reviews factor graphs and their connection to linear algebra and tensor diagrams. Section III makes the pivotal observation that factor graphs with complex factors and with auxiliary variables that are not random variables can represent probability mass functions. The main results are given in Section IV.

We will use standard linear algebra notation rather than the bra-ket notation of quantum mechanics. The Hermitian transpose of a complex matrix A will be denoted by $A^H \triangleq \overline{A^T}$, where A^T is the transpose of A and \overline{A} is the componentwise complex conjugate. An identity matrix will be denoted by I .

II. ON FACTOR GRAPHS AND MATRICES

Factor graphs represent factorizations of functions of several variables. In this paper, all variables take values in finite

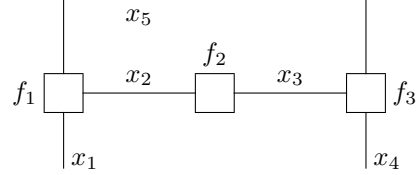


Fig. 1. Forney factor graph of (1).

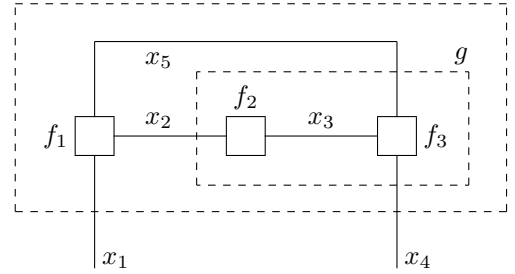


Fig. 2. Closing boxes in factor graphs.

alphabets and all functions take values in \mathbb{C} . We will use Forney factor graphs (also called normal factor graphs) as in [2], [3] where nodes/boxes represent factors and edges represent variables. For example, assume that some function $f(x_1, \dots, x_5)$ can be written as

$$f(x_1, \dots, x_5) = f_1(x_1, x_2, x_5) f_2(x_2, x_3) f_3(x_3, x_4, x_5). \quad (1)$$

The corresponding factor graph is shown in Fig. 1.

The Forney factor-graph notation is intimately connected with the idea of “closing boxes” by summing over internal variables [2]. For example, closing the inner dashed box in Fig. 2 replaces the two nodes/factors $f_2(x_2, x_3)$ and $f_3(x_3, x_4, x_5)$ by the single node/factor

$$g(x_2, x_4, x_5) \triangleq \sum_{x_3} f_2(x_2, x_3) f_3(x_3, x_4, x_5); \quad (2)$$

closing the outer dashed box in Fig. 2 replaces all nodes/factors in (1) by the single node/factor

$$f(x_1, x_4) \triangleq \sum_{x_2, x_3, x_5} f(x_1, \dots, x_5); \quad (3)$$

and closing first the inner dashed box and then the outer dashed box replaces all nodes/factors in (1) by

$$\sum_{x_2, x_5} f_1(x_1, x_2, x_5) g(x_2, x_4, x_5) = f(x_1, x_4). \quad (4)$$

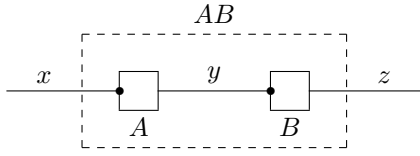


Fig. 3. Factor-graph representation of matrix multiplication (7). The small dot denotes the variable that indexes the rows of the corresponding matrix.

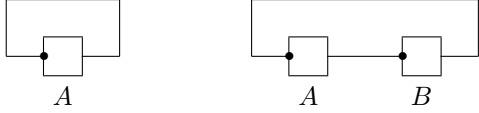


Fig. 4. Factor graph of $\text{tr}(A)$ (left) and of $\text{tr}(AB) = \text{tr}(BA)$ (right).

Note the equality between (4) and (3), which holds in general: closing an inner box within some outer box (by summing over its internal variables) does not change the closed-box function of the outer box.

A *half edge* in a factor graph is an edge that is connected to only one node (such as x_1 in Fig. 1). The *external function* of a factor graph (in [12]–[14] also called *partition function*) is defined to be the closed-box function of a box that contains all nodes and all full edges, but all half edges stick out (such as the outer box in Fig. 2). The external function of Fig. 1 is (3).

The equality constraint function $f_{=}$ is defined as

$$f_{=}(x_1, \dots, x_n) = \begin{cases} 1, & \text{if } x_1 = \dots = x_n \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

The corresponding node (which is denoted by “=”) can serve as a branching point in a factor graph, cf. Figs. 7–11.

A matrix $A \in \mathbb{C}^{m \times n}$ may be viewed as a function

$$\{1, \dots, m\} \times \{1, \dots, n\} \rightarrow \mathbb{C} : (x, y) \mapsto A(x, y). \quad (6)$$

The multiplication of two matrices A and B can then be written as

$$(AB)(x, z) = \sum_y A(x, y)B(y, z), \quad (7)$$

which is the closed-box function (the external function) of Fig. 3. Note that the identity matrix corresponds to an equality constraint function $f_{=}(x, y)$.

In this notation, the trace of a square matrix A is

$$\text{tr}(A) = \sum_x A(x, x), \quad (8)$$

which is the external function (which is a constant) of the factor graph in Fig. 4. Also shown in Fig. 4 is the graphical proof of the identity $\text{tr}(AB) = \text{tr}(BA)$, which is much used in quantum mechanics.

Factor graphs for linear algebra operations such as Fig. 3 and Fig. 4 (and the corresponding generalizations to tensors) are essentially tensor diagrams (or trace diagrams) as in [10], [11]. This connection between factor graphs and tensor diagrams was noted in [12]–[14].

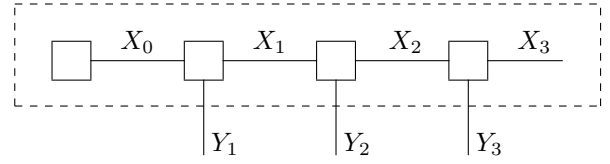


Fig. 5. Factor graph of the hidden Markov model (9) for $n = 3$.

III. STATISTICAL MODELS WITH AUXILIARY VARIABLES USING COMPLEX FACTORS

Statistical models usually contain auxiliary variables in addition to the observable variables. Consider, for example, a hidden Markov model with observables Y_1, \dots, Y_n and auxiliary variables (hidden variables) X_0, X_1, \dots, X_n such that

$$p(y_1, \dots, y_n, x_0, \dots, x_n) = p(x_0) \prod_{k=1}^n p(y_k, x_k | x_{k-1}). \quad (9)$$

The factor graph of (9) is given in Fig. 5. (As shown in this example, variables in factor graphs are often denoted by capital letters [2].) Closing the dashed box in Fig. 5 yields $p(y_1, \dots, y_n)$, the probability mass function of the observables.

As illustrated by this example, auxiliary variables in statistical models are often introduced in order to obtain nice state-space models.

In traditional statistical models, such auxiliary state variables are themselves random variables, and the total model is a joint probability law over all variables as, e.g., in (9). (A statistical model may also contain parameters in addition to auxiliary random variables, but such parameters are not relevant for the present discussion.)

The first main point of this paper is this: *requiring the auxiliary state variables to be random variables may be unnecessarily restrictive*. The benefits of state-space representations may be obtained by merely requiring a function $f(y, x)$ (with a useful factorization) such that the probability mass function of the observables is

$$p(y) = \sum_x f(y, x); \quad (10)$$

the function $f(y, x)$ need not be a probability mass function and it need not even be real valued.

For example, consider the factor graph in Fig. 6, where all factors are complex valued. Note that the lower dashed box in Fig. 6 mirrors the upper dashed box, but all factors in the lower box are the complex conjugates of the corresponding factors in the upper dashed box. The closed-box function of the upper dashed box is

$$g(y_1, y_2, y_3) \triangleq \sum_{x_1, x_2} g_1(x_1, y_1) g_2(x_1, x_2, y_2) g_3(x_2, y_3) \quad (11)$$

and the closed-box function of the lower dashed box is

$$\sum_{x'_1, x'_2} \overline{g_1(x'_1, y_1)} \overline{g_2(x'_1, x'_2, y_2)} \overline{g_3(x'_2, y_3)} = \overline{g(y_1, y_2, y_3)}. \quad (12)$$

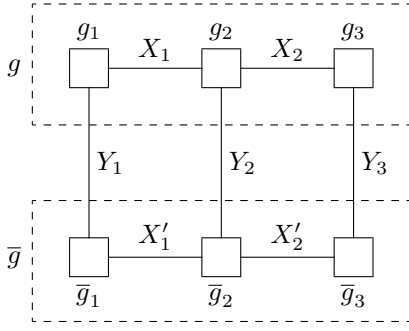


Fig. 6. Factor graph for $p(y_1, y_2, y_3)$ with complex-valued factors.

It follows that the closed-box function in Fig. 6 (with both dashed boxes closed) is

$$g(y_1, y_2, y_3) \overline{g(y_1, y_2, y_3)} = |g(y_1, y_2, y_3)|^2, \quad (13)$$

which is real and nonnegative and thus suitable to represent a probability mass function $p(y_1, y_2, y_3)$ (up to a scale factor).

We will see that factor graphs as in Fig. 6—with two parts, one part being the complex conjugate mirror image of the other part—can represent probabilities in quantum mechanics.

IV. FACTOR GRAPHS FOR QUANTUM MECHANICS

Consider the factor graph of Fig. 7. In this figure, U_0 and U_1 are $M \times M$ unitary matrices, and all variables except Y_1 and Y_2 take values in the set $\{1, \dots, M\}$. The two large boxes in the figure represent measurements, as will be detailed below. The factor/box $p(x_0)$ is a probability mass function over the initial state X_0 . We will see that this factor graph (with suitable modeling of the measurements) represents the joint probability mass function $p(y_1, y_2)$ of a general M -dimensional quantum system with two observations Y_1 and Y_2 . The generalization to more observed variables Y_1, Y_2, \dots is obvious.

The unitary matrices U_0 and U_1 in Fig. 7 represent the development of the system between the initial state and the first measurement, or between measurements, respectively, according to the Schrödinger equation.

In the most basic case, the initial state $X_0 = x_0$ is known and the measurements look as shown in Fig. 8, where the matrices B_1 and B_2 are also unitary. In this case, the observed variables Y_1 and Y_2 take values in $\{1, \dots, M\}$ as well. Note that the lower part of this factor graph is the complex conjugate mirror of the upper part (as in Fig. 6).

In quantum-mechanical terms, measurements as in Fig. 8 are projection measurements with one-dimensional eigenspaces. Note that the value of Y_1 and Y_2 is the index of the measured eigenspace (rather than the corresponding eigenvalue).

A very general form of measurement is shown in Fig. 11. In this case, the range of Y_k is a finite set \mathcal{Y}_k , and for each $y \in \mathcal{Y}_k$, the factor $A_k(\tilde{x}_k, x_k, y)$ corresponds to a complex square matrix $A_k(y)$ (with row index \tilde{x}_k and column index x_k) such that

$$\sum_{y \in \mathcal{Y}_k} A_k(y)^H A_k(y) = I \quad (14)$$

(see [9, Chap. 2]). Measurements as in Fig. 8 are included as a special case with $\mathcal{Y}_k = \{1, \dots, M\}$ and

$$A_k(y) = B_k(y) B_k(y)^H, \quad (15)$$

where $B_k(y)$ denotes the y -th column of B_k .

It is clear from Section III that the external function of Fig. 7 (with measurements as in Fig. 8 or as in Fig. 11) is real and nonnegative. We now proceed to analyze these factor graphs and to verify that they yield the correct quantum-mechanical probabilities $p(y_1, y_2)$ for the respective class of measurements. To this end, we need to understand the closed-box functions of the dashed boxes in Fig. 9. We begin with the dashed box on the right-hand side of Fig. 9, where Y_2 is assumed to be unknown.

Proposition 1. Closing the dashed box on the right-hand side in Fig. 9 (with a measurement of Y_2 as in Fig. 8 or as in Fig. 11, but with unknown result of the measurement) reduces it to an equality constraint function. \square

The proof of this proposition and the proofs of the subsequent propositions are easy, and are omitted due to space constraints.

Proposition 1 guarantees, in particular, that a future measurement (with as yet unknown results) does not influence present or past observations.

The proposition clearly holds also for the extension of Fig. 7 to any finite number of measurements Y_1, Y_2, \dots and can then be applied recursively from right to left.

Applying reductions according to Proposition 1 recursively from right to left in Fig. 7 leads to the following proposition.

Proposition 2. The factor graph of Fig. 7 (with measurements as in Fig. 8 or as in Fig. 11) represents a properly normalized probability mass function, i.e., the external function $p(y_1, y_2)$ is real and $\sum_{y_1, y_2} p(y_1, y_2) = 1$. \square

Consider now the dashed box on the left in Figs. 9 and 10, which turns out to be the density matrix of quantum mechanics. We will distinguish between the closed-box function $\rho_k(x_k, x'_k)$ before measuring Y_k (as in Fig. 9), and the closed-box function $\tilde{\rho}_k(\tilde{x}_k, \tilde{x}'_k)$ after the observation $Y_k = y_k$ (as in Fig. 10). The former is easily seen to be properly normalized, but the latter needs normalization to satisfy (16). The corresponding matrices will be denoted by ρ_k and $\tilde{\rho}_k$, respectively. The proper normalization can then be expressed by the condition

$$\text{tr}(\rho_k) = \text{tr}(\tilde{\rho}_k) = 1. \quad (16)$$

Proposition 3 (Unitary Evolution Between Measurements). The matrix ρ_{k+1} is obtained from the matrix $\tilde{\rho}_k$ as

$$\rho_{k+1} = U_k \tilde{\rho}_k U_k^H. \quad (17)$$

\square

Proposition 4 (Basic Projection Measurement). In Fig. 7 (generalized to any number of observations), if Y_k is measured

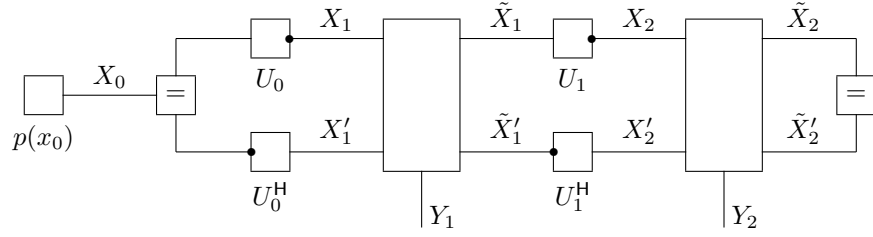


Fig. 7. Factor graph of quantum system with two measurements and the corresponding observations Y_1 and Y_2 .

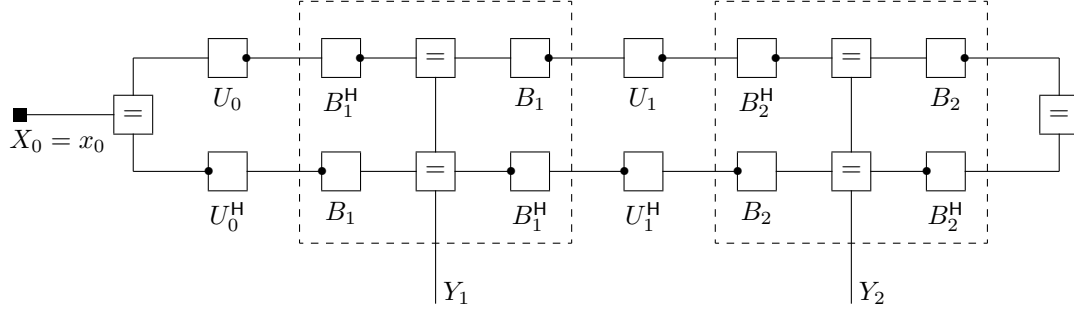


Fig. 8. Important special case of Fig. 7: all matrices are unitary and the initial state $X_0 = x_0$ is known. In quantum-mechanical terms, such measurements are projection measurements with one-dimensional eigenspaces.

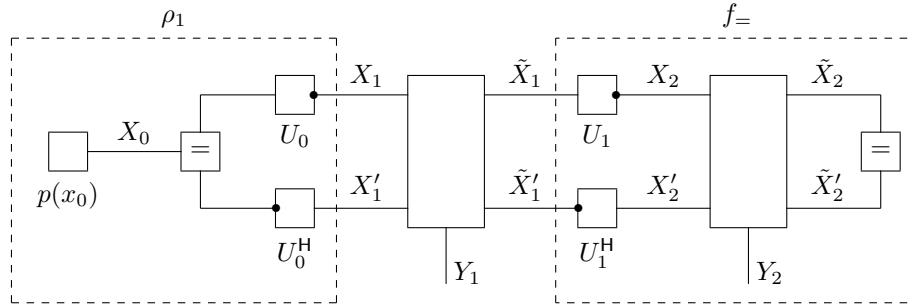


Fig. 9. The closed-box function of the dashed box on the left is the density matrix $\rho_1(x_1, x'_1)$. If Y_2 is not known, the dashed box on the right reduces to the constraint $\tilde{X}_1 = \tilde{X}'_1$.

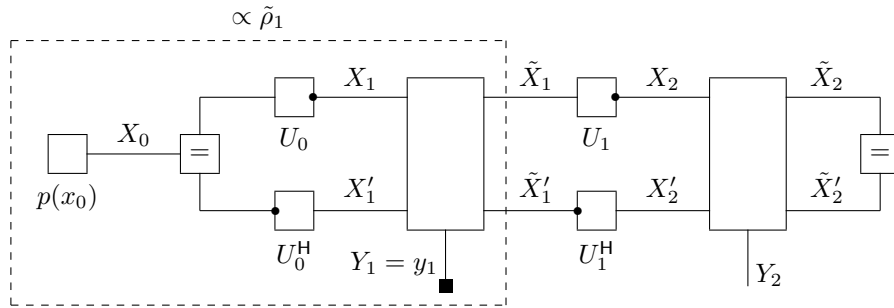


Fig. 10. Density matrix $\tilde{\rho}_1$ after measuring $Y_1 = y_1$.

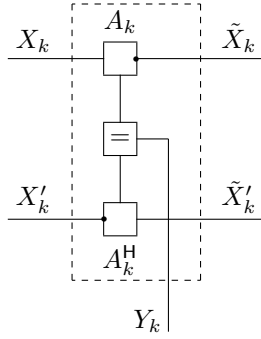


Fig. 11. General measurement as in [9, Chap. 2]. Condition (14) must be satisfied.

as in Fig. 8, then

$$\begin{aligned} P(Y_k = y \mid Y_{k-1} = y_{k-1}, \dots, Y_1 = y_1) \\ = B_k(y)^H \rho_k B_k(y) \end{aligned} \quad (18)$$

$$= \text{tr}(B_k(y) B_k(y)^H \rho_k), \quad (19)$$

where $B_k(y)$ is the y -th column of B_k . After measuring/observing $Y_k = y$, the density matrix is

$$\tilde{\rho}_k = B_k(y) B_k(y)^H. \quad (20)$$

□

Note that (20) is properly normalized because $\text{tr}(B_k(y) B_k(y)^H) = \text{tr}(B_k(y)^H B_k(y)) = \|B_k(y)\|^2 = 1$.

In the special case of Fig. 8, with known initial state $X_0 = x_0$, the matrix ρ_k factors as

$$\rho_k(x_k, x'_k) = \psi_k(x_k) \overline{\psi_k(x'_k)}, \quad (21)$$

or, in matrix notation,

$$\rho_k = \psi_k \psi_k^H, \quad (22)$$

where ψ_k is a column vector of norm 1. The post-measurement density matrix $\tilde{\rho}_k$ factors analogously, as is obvious from (20). In quantum-mechanical terms, ψ_k is the quantum state. The probability (18) can then be expressed as

$$\begin{aligned} P(Y_k = y \mid Y_{k-1} = y_{k-1}, \dots, Y_1 = y_1) \\ = B_k(y)^H \psi_k \psi_k^H B_k(y) \end{aligned} \quad (23)$$

$$= \|B_k(y)^H \psi_k\|^2, \quad (24)$$

which is the most basic form of computing probabilities in quantum mechanics.

Proposition 5 (General Measurement). In Fig. 7 (generalized to any number of observations), if Y_k is measured as in Fig. 11, then

$$\begin{aligned} P(Y_k = y \mid Y_{k-1} = y_{k-1}, \dots, Y_1 = y_1) \\ = \text{tr}(A_k(y) \rho_k A_k(y)^H). \end{aligned} \quad (25)$$

After measuring/observing $Y_k = y$, the density matrix is

$$\tilde{\rho}_k = \frac{A_k(y) \rho_k A_k(y)^H}{\text{tr}(A_k(y) \rho_k A_k(y)^H)} \quad (26)$$

□

According to Propositions 2–5, the factor graph of Fig. 7 (with measurements as in Fig. 8 or as in Fig. 11) yields indeed the correct quantum-mechanical probabilities for the respective class of measurements.

V. CONCLUSION

We have proposed a class of factor graphs that represent quantum-mechanical probabilities involving any number of measurements, both for basic projection measurements and for general measurements as in [9, Chap. 2]. Such factor graphs have not previously been used in statistical modeling.

The space constraints of this paper preclude the discussion of further pertinent topics that we intend to address elsewhere, including the meaning of such factor graphs from a statistical-modeling point of view (disregarding physics), the relation to quantum Bayesian networks (see, e.g., [15]), to quantum belief propagation (see, e.g., [16]), and to tensor diagrams/networks for analyzing quantum systems (see, e.g., [17]). It is also noteworthy that quantum circuits as in [9, Chap. 4] may be viewed as halves of factor graphs as in this paper, where the missing other half is a complex conjugate mirror image as in Section III.

REFERENCES

- [1] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, “Factor graphs and the sum-product algorithm,” *IEEE Trans. Inf. Theory*, vol. 47, pp. 498–519, Feb. 2001.
- [2] H.-A. Loeliger, “An introduction to factor graphs,” *IEEE Sig. Proc. Mag.*, Jan. 2004, pp. 28–41.
- [3] H.-A. Loeliger, J. Dauwels, Junli Hu, S. Korl, Li Ping, and F. R. Kschischang, “The factor graph approach to model-based signal processing,” *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1295–1322, June 2007.
- [4] M. Mézard and A. Montanari, *Information, Physics, and Computation*. Oxford University Press, 2009.
- [5] M. I. Jordan, “Graphical models,” *Statistical Science*, vol. 19, no. 1, pp. 140–155, 2004.
- [6] Ch. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer Science+Business Media, 2006.
- [7] D. Koller and N. Friedman, *Probabilistic Graphical Models*. Cambridge, MA, MIT Press, 2009.
- [8] G. Auletta, M. Fortunato, and G. Parisi, *Quantum Mechanics*. Cambridge University Press, 2009.
- [9] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
- [10] P. Cvitanović, *Group Theory: Birdtracks, Lie’s, and Exceptional Groups*. Princeton Univ. Press, 2008.
- [11] E. Peterson, “Unshackling linear algebra from linear notation,” arXiv:0910.1362, 2009.
- [12] A. Al-Bashabsheh and Y. Mao, “Normal factor graphs and holographic transformations,” *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 752–763, Feb. 2011.
- [13] G. D. Forney, Jr., and P. O. Vontobel, “Partition functions of normal factor graphs,” *Proc. Inf. Theory & Appl. Workshop*, UC San Diego, La Jolla, CA, USA, Feb. 6–11, 2011.
- [14] A. Al-Bashabsheh, Y. Mao, and P. O. Vontobel, “Normal factor graphs: a diagrammatic approach to linear algebra,” *Proc. IEEE Int. Symp. Inf. Theory*, St. Petersburg, Russia, Jul. 31–Aug. 5, 2011, pp. 2178–2182.
- [15] R. R. Tucci, “Quantum information theory – a quantum Bayesian net perspective,” arXiv:quant-ph/9909039v1, 1999.
- [16] M. S. Leifer and D. Poulin, “Quantum graphical models and belief propagation,” *Annals of Physics*, vol. 323, no. 8, pp. 1899–1946, Aug. 2008.
- [17] Z.-C. Gu, M. Levin, and X.-G. Wen, “Tensor-entanglement renormalization group approach as a unified method for symmetry breaking and topological phase transitions,” *Phys. Rev. B*, vol. 78, p. 205116, Nov. 2008.